lemma, using $\vec{X}^i \triangleq (X_1^i, \cdots, X_n^i)$,

$$P_f^n\left(\vec{X}^i = (1,\cdots,1) \text{ infinitely often}\right) = 1,$$

since $\sum_{i=1}^{\infty} P_f^n(\vec{X}^i = (1,\cdots,1)) \geq \sum_{i=1}^{\infty} p_i^n = \infty$. Thus, for any η, one may find a k large enough such that $P_f^n(|J^{\vec{X}}| \leq \eta)$ is arbitrarily small. □

Remark: Note that we have actually shown that, for any fixed n and any $\epsilon < 1 - 1/\log_2(3)$, one may construct a $\mathscr{P}$ and a $\mathscr{C}$ such that the probability of error is arbitrarily close to 1. By defining p_i, $i \geq 2$, to be smaller, we could also take any $\epsilon < 1$.

References

[1] L. G. Valiant, "A theory of the learnable," *Commun. ACM*, vol. 27, no. 11, pp. 1134–1142, 1984.

[2] A. Blumer, A. Ehrenfeucht, D. Haussler, and M. Warmuth, "Learnability and the Vapnik-Chervonenkis dimension," *J. ACM*, vol. 36, no. 4, pp. 929–965, 1989.

[3] D. Haussler, "Decision theoretic generalizations of the PAC model for neural net and other learning applications," *Inf. Comput.*, vol. 20, pp. 78–150, 1992.

[4] G. M. Benedek and A. Itai, "Learnability with respect to a fixed distribution," *Theor. Comput. Sci.*, vol. 86, pp. 377–389, 1991.

[5] V. N. Vapnik, *Estimation of Dependences Based on Empirical Data.* New York: Springer-Verlag, 1982.

[6] R. M. Dudley, "A course on empirical processes," *Lecture Notes in Math. Vol. 1097.* New York: Springer, 1984, pp. 1–142.

[7] V. N. Vapnik and A. Ya. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities," *Theory Probab. Its Appl.*, vol. 16, no. 2, pp. 264–280, 1971.

[8] V. N. Vapnik and A. Ya. Chervonenkis, "Necessary and sufficient conditions for the uniform convergence of means to their expectations," *Theory Probab. Its Appl.*, vol. 26, no. 3, pp. 532–553, 1981.

[9] S. R. Kulkarni, "Problems of computational and information complexity in machine vision and learning," Ph.D. thesis, Dep. Elec. Eng. Comput. Sci., M.I.T., June 1991.

Non White Gaussian Multiple Access Channels with Feedback

Sandeep Pombra and Thomas M. Cover

Abstract—**Although feedback does not increase capacity of an additive white noise Gaussian channel, it enables prediction of the noise for non-white additive Gaussian noise channels and results in an improvement of capacity, but at most by a factor of 2 (Pinsker, Ebert, Pombra, and Cover). Although the capacity of white noise channels cannot be increased by feedback, multiple access white noise channels have a capacity increase due to the cooperation induced by feedback. Thomas has shown that the total capacity (sum of the rates of all the senders) of an m-user Gaussian white noise multiple access channel with feedback is less than twice the total capacity without feedback. In this paper, we show that this factor of 2 bound holds even when cooperation and prediction are combined, by proving that feedback increases the total capacity of an m-user multiple access channel with non-white additive Gaussian noise by at most a factor of 2.**

Manuscript received Oct. 1, 1991; revised Sept. 1, 1993. This work was partially supported by NSF Grant NCR-89-14538, DARPA Contract J-FBI-91-218, and JSEP Contract DAAL03-91-C-0010.

S. Pombra is at 413 Mountain Laurel Court, Mountain View, CA 94043.

T. M. Cover is with the Departments of Electrical Engineering and Statistics, Stanford University, Stanford, CA 94305.

IEEE Log Number 9401815.

Index Terms—**Feedback capacity, Capacity, Multiple-access channel, Non-white Gaussian noise, Gaussian channel.**

I. Introduction

In satellite communication, many senders communicate with a single receiver. The noise in such multiple access channels can often be characterized by non-white additive Gaussian noise. For example, microwave communication components often introduce non-white noise into a channel.

In single-user Gaussian channels with non-white noise, feedback increases capacity. The reason is due solely to the fact that the transmitter knows the past noise (by subtracting out the feedback) and thus can predict the future noise and use this information to increase capacity. A factor of 2 bound on the increase in capacity due to feedback of a single-user Gaussian channel with non-white noise was obtained in [1], [2], [10]. Ihara [9] has shown that the factor of 2 bound is achievable for certain autoregressive additive Gaussian noise channels.

Unlike the simple discrete memoryless channel, feedback in the multiple access channel can increase capacity even when the channel is memoryless, because feedback enables the senders to cooperate with each other. This cooperation is impossible without feedback. This was first demonstrated by Gaarder and Wolf [5]. Cover and Leung [6] established an achievable rate region for the multiple access channel with feedback. Later, Willems [7] proved that the Cover–Leung region is indeed the capacity region for a certain class of channels including the binary adder channel. Ozarow [8] found the capacity region for the two-user Gaussian multiple access channel using a modification of the Kailath–Schalkwijk [4] scheme for simple Gaussian channels. Thomas [11] proved a factor of 2 bound on the capacity increase with feedback for a Gaussian white noise multiple access channel. Keilers [3] characterized the capacity region for a non-white Gaussian noise multiple access channel without feedback. Coding theorems for multiple access channels with finite memory noise are treated in Verdú [14].

The case of non-white Gaussian multiple access channel with feedback combines the above two problems. Here feedback helps through cooperation of senders, as well as through prediction of noise. If we simply use the factor of 2 bounds derived by Cover and Pombra [10] and Thomas [11] for the single-user Gaussian channel with non-white noise and the Gaussian multiple-access channel with white noise, respectively, we might expect feedback to quadruple the total capacity of a non-white m-user Gaussian multiple access channel. However this reasoning is misleading due to the following reasons: Prediction of noise by the receiver and cooperation between the senders are not mutually exclusive events. Also the factor of 2 bound on the feedback capacity of a non-white Gaussian channel has been shown to be tight for the case of only one sender, where there is no interference among the senders. If we have more than one sender, the interference among the senders may diminish the feedback capacity gain due to the prediction of noise.

In this paper, we establish a factor of 2 bound on the increase in total capacity due to feedback for an m-user additive Gaussian non-white noise multiple access channel. Throughout this paper, we define the total capacity of the multiple access channel to be the maximum achievable sum of rates of all the senders.

The paper is organized as follows. In Section II (Theorem 2.1), we prove an expression for the total capacity C_n in bits per transmission of a Gaussian non-white noise multiple access channel without feedback for n uses of the channel. The total capacity C_n is achieved by water filling the total power on the eigenvalues of the noise covariance.

In Section III (Theorem 3.1), an outer bound for the capacity region of a general additive non-white noise multiple access channel with feedback is proved. For Gaussian noise the capacity region is bounded by determinants of the covariances of the inputs and the noise process. We then formally define an upper bound $\overline{C}_{n,FB}$ on the total capacity of a non-white Gaussian multiple access channel with feedback for block length n.

In Section IV, we use this upper bound $\overline{C}_{n,FB}$ to show that feedback increases total capacity of an m-user additive Gaussian non-white noise multiple access channel by at most $\frac{1}{2}\log(m+1)$ bits per transmission.

In Section V, we prove the factor of 2 bound. We do this as follows. First we prove the necessary lemmas in Section V-A. We then define a function $\hat{C}_{n,FB}(K_U)$ whose maximum is $\overline{C}_{n,FB}$. Here K_U is the joint covariance of the inputs and the noise process. In Lemma 5.2 we show that $\hat{C}_{n,FB}(K_U)$ is a concave function of K_U. In Lemma 5.3 we show that $\overline{C}_{n,FB}(P_1, P_2, \cdots, P_m)$ is a concave function of $(P_1, P_2, \cdots, P_m)$. In Section V-B (Theorem 5.1), we prove $\overline{C}_{n,FB} \le 2C_n$ for an even number of equal power senders. First we use the concavity of $\hat{C}_{n,FB}$ (Lemma 5.2) to show that the covariance matrix K_U that maximizes $\overline{C}_{n,FB}$ has a special Toeplitz-like form ((28)). Now by considering the sum of the rates over a subset S of size $m/2$ of the senders, we bound $\overline{C}_{n,FB}$ by $2C_n + T_1$. This bound represents the increase due to cooperation. Then, by considering the sum of the rates over all the senders, we bound $\overline{C}_{n,FB}$ by $2C_n + T_2$. This bound represents the increase due to prediction. Since it can be shown that the terms T_1 and T_2 always have opposite signs, we are able to combine these two bounds to obtain $\overline{C}_{n,FB} \le 2C_n$. In Section V-C, we prove $\overline{C}_{n,FB} \le 2C_n$ for an even number of senders with unequal powers. This is done by using the Schur concavity of $\overline{C}_{n,FB}(P_1, P_2, \cdots, P_m)$ (Lemma 5.3). Here we also use the special properties of C_n (Corollary 2.1 of Section II). Finally in Section V-D, we prove $\overline{C}_{n,FB} \le 2C_n$ for an odd number of senders, with unequal power. Thus we have $\overline{C}_{n,FB} \le 2C_n$ for all cases and for all n. Thus we have proved that feedback at most doubles the total capacity of an m-user additive Gaussian non-white noise multiple access channel.

II. Capacity Without Feedback

Consider an m-user multiple access channel with senders $X_1, X_2, \cdots, X_m$ all sending to the same receiver Y. As shown in Fig. 1 the received signal Y_i at the time i is given by

$$Y_i = \sum_{j=1}^{m} X_{ij} + Z_i, \tag{1}$$

where $Z_1, Z_2, Z_3, \cdots, Z_n$ is a non-white additive Gaussian noise process and X_{ij} denotes the signal sent by sender j at time i. Let $Z^n = (Z_1, \cdots, Z_n) \sim N_n(0, K_Z^{(n)})$. There is a power constraint P_j on each of the senders, i.e., for all senders $j = 1, 2, \cdots, m$, we must have

$$\frac{1}{n}\sum_{i=1}^{n} x_{ij}^2(W_j) \le P_j, \qquad W_j \in \{1, 2, \cdots, 2^{nR_j}\}. \tag{2}$$

The capacity of the non-white noise Gaussian channel without feedback was first characterized by Keilers [3]. In the m-user case the capacity region for n uses of the channel is the set of all

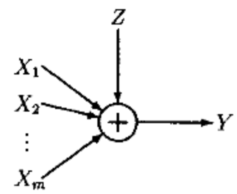

Fig. 1. Gaussian Multiple Access Channel.

rate vectors $(R_1, R_2, \cdots, R_m)$ satisfying

$$\sum_{j\in S} R_j \le \frac{1}{2n}\log\frac{\left|\sum_{j\in S} K_{X_j}^{(n)} + K_Z^{(n)}\right|}{|K_Z^{(n)}|}, \tag{3}$$

for every subset S of the senders $\{1, 2, \cdots, m\}$, for some $n \times n$ covariances $K_{X_j}^{(n)}$ of the vectors $X_j^n = (X_1, \cdots, X_n)$, satisfying the power constraint

$$\frac{1}{n}\operatorname{tr}\left(K_{X_j}^{(n)}\right) \le P_j, \qquad j = 1, 2, \cdots, m.$$

Keilers [3] used a sequential water filling procedure to obtain the extreme points of the convex hull of the capacity region.

We now state a theorem characterizing the total capacity $C_n(P_1, P_2, \cdots, P_m)$ (maximum achievable sum of rates of all the senders) of a Gaussian non-white noise multiple access channel without feedback. The theorem states that $C_n(P_1, P_2, \cdots, P_m)$ is obtained by water filling the total power $\sum_{j=1}^{m} P_j$ on the eigenvalues of the noise covariance. This theorem may be interpreted as follows. The Gaussian multiple access channel represented by noise covariance $K_Z^{(n)}$ is equivalent to n parallel additive white Gaussian noise (AWGN) multiple access channels with noise power given by the eigenvalues of $K_Z^{(n)}$. In each of these AWGN multiple access channels the capacity is solely a function of the average total power of the senders over that channel. Hence the total capacity C_n may be reduced to water filling the total power $\sum_{j=1}^{m} P_j$ On the eigenvalues of $K_Z^{(n)}$.

Theorem 2.1: The total capacity in bits per transmission for n uses of the additive non-white Gaussian multiple access channel without feedback is

$$C_n(P_1, P_2, \cdots, P_m) = \max_{\frac{1}{n}\operatorname{tr}(K_{X_j}^{(n)}) \le P_j,\, j=1,2,\cdots,m} \frac{1}{2n} \log\frac{\left|\sum_{j=1}^{m} K_{X_j}^{(n)} + K_Z^{(n)}\right|}{|K_Z^{(n)}|}. \tag{4}$$

This reduces to water filling the total power $\sum_{j=1}^{m} P_j$ on the eigenvalues $\{\lambda_i^{(n)}\}$ of $K_Z^{(n)}$. Thus

$$C_n(P_1, P_2, \cdots, P_m) = C_n\left(\sum_{j=1}^{m} P_j\right) \tag{5}$$

$$= \frac{1}{2n}\sum_{i=1}^{n}\log\left(1 + \frac{(\lambda - \lambda_i^{(n)})^+}{\lambda_i^{(n)}}\right), \tag{6}$$

where $(y)^+ = \max\{y, 0\}$ and where λ is chosen so that

$$\sum_{i=1}^{n}(\lambda - \lambda_i)^+ = n\sum_{j=1}^{m} P_j. \tag{7}$$

This theorem can be proved using simple linear algebra and the Kuhn–Tucker conditions and is essentially the well-known water filling argument (see [13]).

Theorem 2.1 implies that the total capacity of a non-white Gaussian multiple access channel without feedback is a function solely of the sum of the powers $\Sigma_{j=1}^m P_j$. In particular, this total capacity is the same regardless of whether the total power $\Sigma_{j=1}^m P_j$ is shared equally between the senders or not. We state this result as a corollary to Theorem 2.1, and we will use this to derive the factor of 2 bound.

Corollary 2.1: The total capacity in bits per transmission for n uses of the additive non-white Gaussian multiple access channel without feedback is

$$C_n(P_1, P_2, \cdots, P_m) = C_n(P, P, \cdots, P) = \max_{\frac{1}{n}\operatorname{tr}(K_X^{(n)}) \le P} \frac{1}{2n} \log \frac{|mK_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|}, \quad (8)$$

where $P = 1/m\Sigma_{j=1}^m P_j$.

In [10] the feedback capacity region for the single-user channel was characterized using an asymptotic equipartition argument for a nonergodic Gaussian process. The capacity $C_{n,FB}$ in bits per transmission for n uses of the channel of the time-varying Gaussian channel with feedback is

$$C_{n,FB} = \max_{\frac{1}{n}\operatorname{tr}(K_X^{(n)}) \le P} \frac{1}{2n} \log \frac{|K_{X+Z}^{(n)}|}{|K_Z^{(n)}|}, \quad (9)$$

where the maximization is taken over all X^n of the form

$$X_i = \sum_{j=1}^{i-1} b_{ij} Z_j + V_i, \qquad i = 1, 2, \cdots, n, \quad (10)$$

and V^n is independent of Z^n.

In the next section we prove an outer bound for the capacity region of a general additive non-white noise multiple access channel with feedback. We use this outer bound to define an upper bound on the total capacity in bits per transmission of a non-white Gaussian multiple access channel with feedback.

III. Outer Bound for the Capacity Region with Feedback

We wish to characterize the capacity of time-varying additive Gaussian noise multiple access channels with feedback. At the same time we wish to show that the feedback total capacity bound $\overline{C}_{n,FB}$ and the nonfeedback total capacity C_n obey the inequality $\overline{C}_{n,FB} \le 2C_n$ in bits per transmission. We shall accomplish this by proving an outer bound for the capacity region of a general additive non-white noise multiple access channel with feedback. We specialize this result to the Gaussian case to obtain the desired bound.

To simplify notation, let S denote an arbitrary subset of $\{1, 2, \cdots, m\}$. Let $X_j^i = (X_{1j}, \cdots, X_{ij})$, with the interpretation that X_{ij} is the signal sent by sender j at time i, and X_j^i is the sequence of the first i transmissions of sender j. Let $X^i(S)$ denote the set $\{X_j^i: j \in S\}$ (e.g., if $S = \{1, 3\}$, then $X^i(S) = \{X_1^i, X_3^i\}$). Let $W_1, W_2, \cdots, W_m$ denote the input messages, where each W_j is uniformly distributed in $(1, 2^{nR_j})$ and is independent of the other messages. Let $W(S)$ denote the set $\{W_j: j \in S\}$. Let $\overline{S}$ denote the complement of the set S. The channel, which satisfies $Y_i = \Sigma_{j \in S} X_{ij} + \Sigma_{j \in \overline{S}} X_{ij} + Z_i$, $i = 1, 2, \cdots, n$ has non-white additive noise $Z_1, Z_2, Z_3, \cdots, Z_n$, where $Z^n = (Z_1, \cdots, Z_n) \sim N_n(0, K_Z^{(n)})$ for the Gaussian case. The output is given by $Y^n = \Sigma_{j \in S} X_j^n + \Sigma_{j \in \overline{S}} X_j^n + Z^n$. Since we have feedback, the input symbol X_{ij} of sender j at time i is a function of the message W_j and the past values of the output, i.e., $Y_1, Y_2, \cdots, Y_{i-1}$. Thus $X_{ij} = X_{ij}(W_j, Y^{i-1})$. For block length n we must specify a

$$((2^{nR_1}, 2^{nR_2}, \cdots, 2^{nR_m}), n)$$

code with codewords

$$x_j^n(W_j, Y^{n-1}) = (x_{1j}(W_j), x_{2j}(W_j, Y^1), \cdots, x_{nj}(W_j, Y^{n-1})),$$
$$W_j \in \{1, 2, \cdots, 2^{nR_j}\} \quad \text{for all } j = 1, 2, \cdots, m.$$

In addition, we require that the codewords satisfy the expected power constraints

$$E\left[\frac{1}{n}\sum_{i=1}^n x_{ij}^2(W_j, Y^{i-1})\right] \le P_j, \qquad j = 1, 2, \cdots, m, \quad (11)$$

where the expectation is taken over all possible noise sequences.

Before we proceed to the outer bound, we need a simple lemma that shows that the output entropy given the inputs $W_1, W_2, \cdots, W_m$ is equal to the entropy of the noise.

Lemma 3.1: For the non-white additive noise multiple access channel with feedback,

$$h(Y^n | W(S), W(\overline{S})) = h(Z^n). \quad (12)$$

This lemma is proved in the Appendix. Lemma 3.1 can easily be strengthened to Lemma 3.2.

Lemma 3.2: For the non-white additive noise multiple access channel with feedback, there exist $n \times n$ lower triangular matrices L_S, $S \subseteq \{1, 2, \cdots, m\}$, such that

$$h(Y^n | W(\overline{S})) = \frac{1}{n}\left(h\left(\sum_{j \in S} X_j^n - L_S \sum_{j \in \overline{S}} X_j^n + Z^n\right) - h(Z^n)\right), \quad (13)$$

for all subsets S of the senders.

Note that this lemma degenerates to Lemma 3.1 when $L_S = 0$ and $\overline{S}$ is the set of all users. Now we are ready to state the outer bound on the capacity region with feedback.

Theorem 3.1: For the non-white additive noise multiple access channel with feedback, a rate vector $(R_1, R_2, \cdots, R_m)$ is achievable only if there exists a feedback code such that

$$\sum_{j \in S} R_j \le \frac{1}{n}\left(h\left(\sum_{j \in S} X_j^n - \sum_{j \in \overline{S}} X_j^n + Z^n\right) - h(Z^n)\right), \quad (14)$$

for all $S \subseteq \{1, 2, \cdots, m\}$, where the joint distribution on $(X_1^n, X_2^n, \cdots, X_m^n, Y^n)$ is induced by the given feedback code. If Z^n is multivariate Gaussian, the rates also satisfy

$$\sum_{j \in S} R_j \le \frac{1}{2n} \log \frac{\left|K^{(n)}_{\Sigma_{j \in S} X_j - \Sigma_{j \in \overline{S}} X_j + Z}\right|}{|K_Z^{(n)}|}. \quad (15)$$

This theorem is proved using Lemma 3.1 and Lemma 3.2 in the Appendix. Thus we get the desired bound on the sum of rates needed to show the factor of 2 bound

$$\sum_{j \in S} R_j \le \frac{1}{n}\left(h\left(\sum_{j \in S} X_j^n - \sum_{j \in \overline{S}} X_j^n + Z^n\right) - h(Z^n)\right), \quad (16)$$

for all subsets S of the senders, where the joint distribution on

$(X_1^n, X_2^n, \cdots, X_m^n, Y^n)$ is induced by the given feedback code. If Z^n is multivariate Gaussian we have

$$\sum_{j \in S} R_j \le \frac{1}{2n} \log \frac{\left| K^{(n)}_{\sum_{j \in S} X_j - \sum_{j \in \bar{S}} X_j + Z} \right|}{|K_Z^{(n)}|}. \tag{17}$$

Thus we can bound the total capacity in terms of determinants of the covariances of the inputs and noise process.

The bounds of (16) and (17) may be interpreted as special cases of more general outer bounds. These general outer bounds can be derived similarly to Theorem 3.1 using Lemma 3.2. We state them below as Theorem 3.2.

Theorem 3.2: For the non-white additive noise multiple access channel with feedback, a rate vector $(R_1, R_2, \cdots, R_m)$ is achievable only if there exists a feedback code such that

$$\sum_{j \in S} R_j \le \frac{1}{n} \left(h\left(\sum_{j \in S} X_j^n - L_S \sum_{j \in \bar{S}} X_j^n + Z^n \right) - h(Z^n) \right), \tag{18}$$

for all subsets S of the senders, for some $n \times n$ lower triangular matrices

$$L_S, \qquad S \subseteq \{1, 2, \cdots, m\},$$

where the joint distribution on $(X_1^n, X_2^n, \cdots, X_m^n, Y^n)$ is induced by the given feedback code. If Z^n is multivariate Gaussian the rates also satisfy

$$\sum_{j \in S} R_j \le \frac{1}{2n} \log \frac{\left| K^{(n)}_{\sum_{j \in S} X_j - L_S \sum_{j \in \bar{S}} X_j + Z} \right|}{|K_Z^{(n)}|}. \tag{19}$$

Note the bounds of (16) and (17) may be obtained with

$$L_S = I^n, \qquad S \subseteq \{1, 2, \cdots, m\}.$$

These outer bounds may not be tight in general. The basic problem in determining feedback multiple access channel capacity is to find the class of joint distributions achievable using feedback. However, for Gaussian multiple access channels the outer bound of (19) in conjunction with the class of joint distributions induced by the form $X_j^n = B_j Y^n + V_j^n$, $j = 1, 2, \cdots, m$, may be used to characterize the capacity region. Here B_j is a strictly lower triangular $n \times n$ matrix, and V_j^n and Z^n are all independent, for $j = 1, 2, \cdots, m$.

Now we use (17) to formally define $\overline{C}_{n,FB}$. Here $\overline{C}_{n,FB}$ is an upper bound on the total capacity in bits per transmission of a non-white Gaussian multiple access channel with feedback if the channel is used for time block $\{1, 2, \cdots, n\}$.

Definition 3.1: For each $n(m+1) \times n(m+1)$ joint covariance K_U of the vector $U = (X_1^n, X_2^n, \cdots, X_m^n, Z^n)^t$ induced by a feedback code, and subject to the power constraints

$$\frac{1}{n} \operatorname{tr}\left(K_{X_j}^{(n)} \right) \le P_j, \qquad j = 1, 2, \cdots, m,$$

there exists a rate region defined by 2^m constraints of (17). Now we consider the region that is the union of these rate regions over the convex hull of all joint covariances K_U. By Theorem 3.1 this region is an outer bound for the capacity region with feedback. Therefore we may define

$$\overline{C}_{n,FB}(P_1, P_2, \cdots, P_m) \triangleq \max \sum_{j=1}^{m} R_j, \tag{20}$$

where the maximization is taken over all rates $(R_1, R_2, \cdots, R_m)$ belonging to the outer bound for the capacity region with feedback defined above.

In the next two sections we use $\overline{C}_{n,FB}$ to establish bounds on the increase in total capacity due to feedback for an m-user additive Gaussian non-white noise multiple access channel.

IV. An Additive Bound

In this section, we use this upper bound $\overline{C}_{n,FB}$ to show that feedback increases total capacity of an m-user additive Gaussian non-white noise multiple access channel by at most $\frac{1}{2}\log(m+1)$ bits per transmission. First we need the following lemma.

Lemma 4.1:

$$\left| K^{(n)}_{(\sum_{j=1}^m X_j)} \right| \le \left| m \sum_{j=1}^{m} K_{X_j}^{(n)} \right| = m^n \left| \sum_{j=1}^{m} K_{X_j}^{(n)} \right|.$$

Proof: By expanding the covariance of the sum of m random vectors we have the following identity

$$K^{(n)}_{(\sum_{j=1}^m X_j)} + \sum_{i=1}^{m} \sum_{j=1}^{i-1} K^{(n)}_{(X_i - X_j)} = m \sum_{j=1}^{m} K_{X_j}^{(n)}. \tag{21}$$

Now $K^{(n)}_{(X_i - X_j)}$, being a covariance matrix, is nonnegative definite symmetric for all i, j. The sum of such nonnegative definite symmetric matrices

$$\sum_{i=1}^{m} \sum_{j=1}^{i-1} K^{(n)}_{(X_i - X_j)}$$

is nonnegative definite symmetric. Therefore we can apply Lemma 1 of [10] to obtain the required result (Lemma (4.1)). □

Theorem 4.1:

$$\overline{C}_{n,FB} \le C_n + \frac{1}{2} \log(m+1).$$

Proof: Let the $n \times n$ covariance matrix $K^{(n)}_{(\sum_{j=1}^m X_j + Z)}$ achieve $\overline{C}_{n,FB}$ in Definition 3.1. Then by using the definition of $\overline{C}_{n,FB}$ and (17) we have

$$\begin{aligned}
\overline{C}_{n,FB} &\le \frac{1}{2n} \log \frac{\left| K^{(n)}_{(\sum_{j=1}^m X_j + Z)} \right|}{|K_Z^{(n)}|} \\
&\le \frac{1}{2n} \log \frac{\left| (m+1)\left(\sum_{j=1}^{m} K_{X_j}^{(n)} + K_Z^{(n)} \right) \right|}{|K_Z^{(n)}|} \\
&= \frac{1}{2n} \log \frac{(m+1)^n \left| \sum_{j=1}^{m} K_{X_j}^{(n)} + K_Z^{(n)} \right|}{|K_Z^{(n)}|} \\
&= \frac{1}{2n} \log \frac{\left| \sum_{j=1}^{m} K_{X_j}^{(n)} + K_Z^{(n)} \right|}{|K_Z^{(n)}|} + \frac{\log(m+1)}{2} \\
&\le C_n + \frac{\log(m+1)}{2}.
\end{aligned} \tag{22}$$

Here, the second inequality, which relates feedback to nonfeedback, follows from Lemma (4.1). The last inequality follows from Theorem 2.1. □

V. Factor of 2 Bound

We will now prove that feedback at most doubles the total capacity of the Gaussian multiple access channel. We use the outer bound of Section III to prove the factor of 2 bound. First we will show some required properties of $\overline{C}_{n,FB}$.

A. Necessary Lemmas

Let $U = (X_1^n, X_2^n, \cdots, X_m^n, Z^n)^t$ be the $n(m+1)$-vector of inputs and noise. Let K_U be the $n(m+1) \times n(m+1)$ covariance matrix of U. We define a function $\hat{C}_{n,FB}(K_U)$ corresponding to $\overline{C}_{n,FB}$ as follows.

Definition 5.1:

$$\hat{C}_{n,FB}(K_U) \triangleq \max \sum_{j=1}^{m} R_j, \tag{23}$$

where the maximization is taken over all rates $(R_1, R_2, \cdots, R_m)$ subject to constraints

$$\sum_{j \in S} R_j \le I_S \quad \text{for all subsets } S \text{ of } \{1, 2, \cdots, m\}, \tag{24}$$

where

$$I_S = \frac{1}{2n} \log \frac{\left| K^{(n)}_{\Sigma_{j \in S} X_j - \Sigma_{j \in \bar{S}} X_j + Z} \right|}{\left| K_Z^{(n)} \right|}. \tag{25}$$

In Lemma 5.2 we show that $\hat{C}_{n,FB}(K_U)$ is a concave function of K_U. In Lemma 5.3 we show that $\overline{C}_{n,FB}(P_1, P_2, \cdots, P_m)$ is a concave function of $(P_1, P_2, \cdots, P_m)$. First we show Lemma 5.1, which states that the maximum of a linear function of an n-vector $\mathscr{R} = (R_1, R_2, \cdots, R_n)^t$ over p linear constraints $\mathscr{I} = (I_1, I_2, \cdots, I_p)^t$ is an increasing and concave function of the constraints $\mathscr{I}$.

Lemma 5.1: Let $V(\mathscr{I})$ be defined as

$$V(\mathscr{I}) \triangleq \max_{A\mathscr{R} \le \mathscr{I}} \mathscr{R}^t b, \tag{26}$$

where A is any $p \times n$ matrix and b is any $n \times 1$ vector. Then $V(\mathscr{I})$ satisfies two properties:

i) *Concavity.* $V(\mathscr{I})$ is a concave function of the constraints $\mathscr{I}$, i.e., for any $p \times 1$ vectors $\mathscr{I}_1, \mathscr{I}_2$, and $0 \le \lambda \le 1$,

$$\lambda V(\mathscr{I}_1) + (1-\lambda) V(\mathscr{I}_2) \le V(\lambda \mathscr{I}_1 + (1-\lambda)\mathscr{I}_2).$$

ii) *Domination.* $V(\mathscr{I})$ is an increasing function of the constraints $\mathscr{I}$, i.e., for any $p \times 1$ vectors $\mathscr{I}_1, \mathscr{I}_2$, if $\mathscr{I}_1$ is component by component less than or equal to $\mathscr{I}_2$, then

$$V(\mathscr{I}_1) \le V(\mathscr{I}_2).$$

The proof is straightforward (see [13]).

Lemma 5.2: Let $\hat{C}_{n,FB}(K_U)$ be as defined in Definition 5.1. Then $\hat{C}_{n,FB}(K_U)$ is a concave function of the matrix K_U, i.e., for any $n(m+1) \times n(m+1)$ nonnegative definite matrices K_{U_1}, K_{U_2}, and $0 \le \lambda \le 1$,

$$\hat{C}_{n,FB}(\lambda K_{U_1} + (1-\lambda) K_{U_2}) \ge \lambda \hat{C}_{n,FB}(K_{U_1}) + (1-\lambda)\hat{C}_{n,FB}(K_{U_2}).$$

The proof easily follows from Lemma 5.1 (see [13]).

Lemma 5.3: Let $\overline{C}_{n,FB}$ be as defined in Definition 3. Let $\mathscr{P} = (P_1, P_2, \cdots, P_m)^t$ be the $m \times 1$ vector of the power constraints. Then $\overline{C}_{n,FB}(\mathscr{P})$ is a concave function of the vector $\mathscr{P}$, i.e., for any $m \times 1$ vectors of power constraints $\mathscr{P}_1, \mathscr{P}_2$, and $0 \le \lambda \le 1$,

$$\overline{C}_{n,FB}(\lambda \mathscr{P}_1 + (1-\lambda)\mathscr{P}_2) \ge \lambda \overline{C}_{n,FB}(\mathscr{P}_1) + (1-\lambda)\overline{C}_{n,FB}(\mathscr{P}_2).$$

The proof easily follows from Lemma 5.2 (see [13]).

Finally we need the following simple lemma which states that information processing from causal feedback does not reduce the entropy.

Lemma 5.4: Let $X_1^n(W_1, Y^{n-1}), X_2^n(W_2, Y^{n-1}), \cdots, X_m^n(W_m, Y^{n-1})$ be a given feedback code and let Z^n be the non-white Gaussian noise vector. Then

$$\left| K_Z^{(n)} \right| \le \left| K^{(n)}_{\Sigma_{j=1}^m X_j - Z} \right|. \tag{27}$$

By the log-concavity of determinants this inequality is also true for all matrices $K^{(n)}_{\Sigma_{j=1}^m X_j - Z}$ in the convex hull of covariances induced by feedback codes.

This lemma easily follows from Lemma 3.2 and the entropy maximization property of the normal distribution (see [10] and [13]).

B. Case of Equal Powers, m Even

We first consider all the powers to be equal, i.e., $P_1 = P_2 = \cdots = P_m = P$, with m even. Let $U = (X_1^n, X_2^n, \cdots, X_m^n, Z^n)^t$ be the vector of inputs and noise. Let K_U be the $n(m+1) \times n(m+1)$ covariance matrix of U. We first show that one of the possible values of K_U that achieves $\overline{C}_{n,FB}$ is the symmetric partially block Toeplitz form

$$K_U = \begin{pmatrix} K_X^{(n)} & K_{X_1X_2}^{(n)} & K_{X_1X_2}^{(n)} & \cdots & K_{X_1X_2}^{(n)} & K_{XZ}^{(n)} \\ K_{X_1X_2}^{(n)} & K_X^{(n)} & K_{X_1X_2}^{(n)} & \cdots & K_{X_1X_2}^{(n)} & K_{XZ}^{(n)} \\ K_{X_1X_2}^{(n)} & K_{X_1X_2}^{(n)} & K_X^{(n)} & \cdots & K_{X_1X_2}^{(n)} & K_{XZ}^{(n)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ K_{X_1X_2}^{(n)} & K_{X_1X_2}^{(n)} & K_{X_1X_2}^{(n)} & \cdots & K_X^{(n)} & K_{XZ}^{(n)} \\ K_{ZX}^{(n)} & K_{ZX}^{(n)} & K_{ZX}^{(n)} & \cdots & K_{ZX}^{(n)} & K_Z^{(n)} \end{pmatrix}, \tag{28}$$

where $K_{X_jX_j}^{(n)} = K_X^{(n)}$ is the covariance of each of the senders $(X_j^n, j = 1, 2, \cdots, m)$, $K_{X_{j_1}X_{j_2}}^{(n)} = K_{X_1X_2}^{(n)}$ is the cross-covariance between any two senders

$$((X_{j_1}^n, X_{j_2}^n), j_1 = 1, 2, \cdots, m, \quad j_2 = 1, 2, \cdots, m),$$

$K_{XZ}^{(n)}, K_{ZX}^{(n)}$ are the cross-covariances of each of the senders $(X_j^n, j = 1, 2, \cdots, m)$ with the noise Z^n, and $K_Z^{(n)}$ is the covariance of the noise. (The following argument uses the symmetrization method given by Thomas [11].)

Let us assume that there is some other form of K that achieves $\overline{C}_{n,FB}$ in Definition 3, i.e.,

$$\overline{C}_{n,FB} = \hat{C}_{n,FB}(K).$$

Since all senders have equal powers, by appropriately relabeling the rows and columns of K, we have a new matrix $\hat{K}$ which also achieves

$$\overline{C}_{n,FB} = \hat{C}_{n,FB}(\hat{K}).$$

By Lemma 5.2, $\hat{C}_{n,FB}(K)$ is a concave function of the matrix K. Therefore the average $\frac{1}{2}(K + \hat{K})$ is a more symmetric form that achieves $\overline{C}_{n,FB}$. Proceeding in this way, by averaging over all $m!$ possible permutations of the rows and columns of K (corresponding to all possible permutations of the senders), we obtain the symmetric form of (28) is given by

$$K_U = \frac{1}{m!} \sum_{\pi \in S_m} K(\pi).$$

Since all senders have equal power P, K_U satisfies the power constraint. Hence, we can restrict our attention to this form of K_U, and we will obtain our bounds using it.

We are now ready to prove $\overline{C}_{n,FB} \le 2C_n$ for the case of equal powers and even m. The outline of the proof is as follows. First by considering the sum of the rates over a subset S of size $m/2$ we bound the total capacity with feedback by $2C_n + T_1$. This is done by using the outer bound of Section III and the symmetric form given in (28). Then by considering the sum of the rates of all the senders we bound the total capacity with feedback by $2C_n + T_2$. This is done by using the outer bound of Section III, Lemma 5.4, the log-concavity of determinants, and the symmetric form in (28). Since it turns out that the terms T_1 and T_2 always have opposite signs, we are able to combine these two bounds to obtain the factor of 2 bound.

Theorem 5.1: For additive Gaussian non-white noise multiple access channels with an even number of equal power senders, we have

$$\overline{C}_{n,FB}(P, P,\cdots, P) \le 2C_n(P, P,\cdots, P).$$

Proof: Consider the symmetric form K_U (28) that achieves $\overline{C}_{n,FB}$ in Definition 3.1. Now consider subsets $S, \overline{S}$ of equal size $m/2$. Then from the symmetric form of (28) we have

$$\begin{aligned} K^{(n)}_{\Sigma_{j\in S}X_j - \Sigma_{j\in \overline{S}}X_j + Z} &= K^{(n)}_{\Sigma_{j\in S}X_j - \Sigma_{j\in \overline{S}}X_j} + K_Z^{(n)} \\ &= 2\left(\frac{m}{2}K_X^{(n)} + \left(\frac{m}{2}\right)\left(\frac{m}{2}-1\right)K^{(n)}_{X_1X_2}\right) \\ &\quad - 2\left(\frac{m}{2}\right)^2 K^{(n)}_{X_1X_2} + K_Z^{(n)} \\ &= mK_X^{(n)} - mK^{(n)}_{X_1X_2} + K_Z^{(n)}. \end{aligned} \tag{29}$$

Now substituting this equation in the constraint $I(S)$ (17), we have

$$\begin{aligned} \sum_{j\in S} R_j &\le \frac{1}{2n}\log\frac{\left|K^{(n)}_{\Sigma_{j\in S}X_j - \Sigma_{j\in \overline{S}}X_j + Z}\right|}{|K_Z^{(n)}|} \\ &= \frac{1}{2n}\log\frac{\left|mK_X^{(n)} - mK^{(n)}_{X_1X_2} + K_Z^{(n)}\right|}{|K_Z^{(n)}|} \\ &\overset{(a)}{=} \frac{1}{2n}\log\frac{|mK_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|} \\ &\quad + \frac{1}{2n}\log\left|I_n - mD^{-1/2}K^{(n)}_{X_1X_2}(D^{-1/2})^t\right| \end{aligned} \tag{30}$$

$$\begin{aligned} &\overset{(b)}{\le} \frac{1}{2n}\log\frac{|mK_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|} \\ &\quad + \frac{1}{2}\log\frac{1}{n}\operatorname{tr}\left[I_n - mD^{-1/2}K^{(n)}_{X_1X_2}(D^{-1/2})^t\right] \\ &\overset{(c)}{=} \frac{1}{2n}\log\frac{|mK_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|} \\ &\quad + \frac{1}{2}\log\left(1 - \frac{m}{n}\operatorname{tr}\left[D^{-1/2}K^{(n)}_{X_1X_2}(D^{-1/2})^t\right]\right), \end{aligned} \tag{31}$$

where $D = mK_X^{(n)} + K_Z^{(n)}$, I_n is the $n \times n$ identity matrix, and $A^{1/2}$ denotes the square root of the nonnegative definite matrix A. Here (a) follows from the fact that D is a nonnegative definite matrix and the identity $|AB| = |A||B|$ for the product of determinants. Inequality (b) is derived from the following argument. Since

$$I_n - mD^{-1/2}K^{(n)}_{X_1X_2}(D^{-1/2})^t = D^{-1/2}K^{(n)}_{\Sigma_{j\in S}X_j - \Sigma_{j\in \overline{S}}X_j + Z}(D^{-1/2})^t$$

is a nonnegative definite matrix, we can use the arithmetic-geometric mean inequality [12] for nonnegative definite matrices:

$$|A|^{1/n} \le \frac{1}{n}\operatorname{tr}(A)$$

on the second term of (30) to obtain (b). The equality (c) follows from $\operatorname{tr}(I_n) = n$, and $\operatorname{tr}(A + B) = \operatorname{tr}(A) + \operatorname{tr}(B)$. By symmetry we can obtain an equation similar to (31) for the sum of the rates over the subset $\overline{S}$ as

$$\begin{aligned} \sum_{j\in \overline{S}} R_j &\le \frac{1}{2n}\log\frac{|mK_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|} \\ &\quad + \frac{1}{2}\log\left(1 - \frac{m}{n}\operatorname{tr}\left[D^{-1/2}K^{(n)}_{X_1X_2}(D^{-1/2})^t\right]\right). \end{aligned} \tag{32}$$

Hence from (31) and (32) we have

$$\sum_{j=1}^{m} R_j \le 2\frac{1}{2n}\log\frac{|mK_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|} + T_1, \tag{33}$$

where

$$T_1 = \log\left(1 - \frac{m}{n}\operatorname{tr}\left[D^{-1/2}K^{(n)}_{X_1X_2}(D^{-1/2})^t\right]\right). \tag{34}$$

Now using Corollary 2.1 and the definition of $\overline{C}_{n,FB}$ in (33), we have

$$\overline{C}_{n,FB}(P, P,\cdots, P) \le 2C_n(P, P,\cdots, P) + T_1. \tag{35}$$

Now we consider the bound on the sum of the rates of all the senders. We have from (17)

$$\begin{aligned} \sum_{j=1}^{m} R_j &\le \frac{1}{2n}\log\frac{\left|K^{(n)}_{\Sigma_{j=1}^{m}X_j + Z}\right|}{|K_Z^{(n)}|} \\ &= 2\frac{1}{2n}\log\frac{\left|K^{(n)}_{\Sigma_{j=1}^{m}X_j + Z}\right|^{1/2}|K_Z^{(n)}|^{1/2}}{|K_Z^{(n)}|} \\ &\overset{(a)}{\le} 2\frac{1}{2n}\log\frac{\left|K^{(n)}_{\Sigma_{j=1}^{m}X_j + Z}\right|^{1/2}\left|K^{(n)}_{\Sigma_{j=1}^{m}X_j - Z}\right|^{1/2}}{|K_Z^{(n)}|} \\ &\overset{(b)}{\le} 2\frac{1}{2n}\log\frac{\left|\frac{1}{2}K^{(n)}_{\Sigma_{j=1}^{m}X_j + Z} + \frac{1}{2}K^{(n)}_{\Sigma_{j=1}^{m}X_j - Z}\right|}{|K_Z^{(n)}|} \\ &\overset{(c)}{=} 2\frac{1}{2n}\log\frac{\left|K^{(n)}_{\Sigma_{j=1}^{m}X_j} + K_Z^{(n)}\right|}{|K_Z^{(n)}|} \\ &\overset{(d)}{=} 2\frac{1}{2n}\log\frac{\left|mK_X^{(n)} + m(m-1)K^{(n)}_{X_1X_2} + K_Z^{(n)}\right|}{|K_Z^{(n)}|} \\ &\overset{(e)}{=} 2\left(\frac{1}{2n}\log\frac{|mK_X^{(n)} + K_Z^{(n)}|}{|K_Z^{(n)}|}\right. \\ &\quad + \frac{1}{2n}\log\left|I_n + m(m-1)\right. \\ &\quad \left.\left.\cdot D^{-1/2}K^{(n)}_{X_1X_2}(D^{-1/2})^t\right|\right) \end{aligned} \tag{36}$$

$$\overset{(f)}{\leq} 2\left(\frac{1}{2n}\log\frac{|mK_X^{(n)}+K_Z^{(n)}|}{|K_Z^{(n)}|} + \frac{1}{2}\log\frac{1}{n}\operatorname{tr}\left[I_n + m(m-1) \cdot D^{-1/2}K_{X_1X_2}^{(n)}(D^{-1/2})'\right]\right)$$

$$\overset{(g)}{=} 2\left(\frac{1}{2n}\log\frac{|mK_X^{(n)}+K_Z^{(n)}|}{|K_Z^{(n)}|} + \frac{1}{2}\log\left(1+\frac{m(m-1)}{n} \cdot \operatorname{tr}\left[D^{-1/2}K_{X_1X_2}^{(n)}(D^{-1/2})'\right]\right)\right). \quad (37)$$

Here (a) follows from Lemma 5.4, (b) follows from the log-concavity of determinants ([10], Lemma 4), (c) is a simple matrix identity ([10], Lemma 2), (d) follows from the symmetric form of K_U, (e) follows from the fact that D is a nonnegative definite matrix and from the identity for the product of determinants $|AB| = |A||B|$. The inequality (f) is derived from the following argument. Since

$$I_n + m(m-1)D^{-1/2}K_{X_1X_2}^{(n)}(D^{-1/2})' = D^{-1/2}\left(K_{\Sigma_{j=1}^m X_j}^{(n)} + K_Z^{(n)}\right)(D^{-1/2})'$$

is a nonnegative definite matrix, we can use the arithmetic-geometric mean inequality for nonnegative definite matrices:

$$|A|^{1/n} \leq \frac{1}{n}\operatorname{tr}(A)$$

on the second term of (36) to obtain (f). The equality (g) follows from $\operatorname{tr}(I_n) = n$, and $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$. Hence from (37) we have

$$\sum_{j=1}^m R_j \leq 2\frac{1}{2n}\log\frac{|mK_X^{(n)}+K_Z^{(n)}|}{|K_Z^{(n)}|} + T_2, \quad (38)$$

where

$$T_2 = \log\left(1+\frac{m(m-1)}{n}\operatorname{tr}\left[D^{-1/2}K_{X_1X_2}^{(n)}(D^{-1/2})'\right]\right). \quad (39)$$

Now using Corollary 2.1 and the definition of $\overline{C}_{n,FB}$ in (38), we have

$$\overline{C}_{n,FB}(P,P,\cdots,P) \leq 2C_n(P,P,\cdots,P) + T_2. \quad (40)$$

We observe from (34) and (39) that

$$\operatorname{sgn}\left(\operatorname{tr}\left[D^{-1/2}K_{X_1X_2}^{(n)}(D^{-1/2})'\right]\right) = -\operatorname{sgn}(T_1) = \operatorname{sgn}(T_2).$$

Hence the terms T_1 and T_2 always have opposite signs, i.e.,

$$\min(T_1, T_2) \leq 0.$$

Therefore we have from (35) and (40),

$$\overline{C}_{n,FB}(P,P,\cdots,P) \leq 2C_n(P,P,\cdots,P) + \min(T_1,T_2) \leq 2C_n(P,P,\cdots,P). \quad \square \quad (41)$$

We have thus proved that the total capacity can at most be doubled using feedback for even m with equal powers.

C. Case of Unequal Powers, m Even

We will use an argument similar to the one used by Thomas [11] to establish the factor of 2 bound from unequal powers with even m. So far we have been dealing only with the case when all the transmitters have the same power constraints. Now let us assume that the transmitter powers are $P_1, P_2, \cdots, P_m$.

1) *Without Feedback*: The dominating constraint on the sum of the rates from Theorem 2.1 is

$$C_n(P_1,P_2,\cdots,P_m) = \frac{1}{2n}\sum_{i=1}^n \log\left(1+\frac{(\lambda-\lambda_i^{(n)})^+}{\lambda_i^{(n)}}\right), \quad (42)$$

where $(y)^+ = \max\{y, 0\}$ and where λ is chosen so that

$$\sum_{i=1}^n (\lambda - \lambda_i)^+ = n\sum_{j=1}^m P_j. \quad (43)$$

Defining $P = 1/m\Sigma_{j=1}^m P_j$, then

$$C_n(P_1,P_2,\cdots,P_m) = C_n(P,P,\cdots,P) = C_n\left(\sum_{j=1}^m P_j\right). \quad (44)$$

2) *With Feedback*: From Definition 3.1 it is easy to see that $\overline{C}_{n,FB}(P_1,P_2,\cdots,P_m)$ is a symmetric function of its arguments $P_1, P_2, \cdots, P_m$. In Lemma 5.3 we showed that $\overline{C}_{n,FB}(\cdot,\cdot,\cdots,\cdot)$ is a concave function. Hence by the properties of symmetric concave functions ([12], p. 104), we have

$$\begin{aligned}\overline{C}_{n,FB}(P_1,P_2,\cdots,P_m) &\leq \overline{C}_{n,FB}\left(\frac{\sum P_j}{m}, \frac{\sum P_j}{m}, \cdots, \frac{\sum P_j}{m}\right)\\ &= \overline{C}_{n,FB}(P,P,\cdots,P)\\ &\leq 2C_n(P,P,\cdots,P)\\ &\qquad\text{(result for equal powers)}\\ &= 2C_n(P_1,P_2,\cdots,P_m). \quad (45)\end{aligned}$$

Hence, even with different powers at the different transmitters, the total capacity with feedback is less than twice the total capacity without feedback for even m.

D. Case of Unequal Powers, m Odd

So far we have been dealing only with the case when m is even. Now let us assume that m is odd. Consider an augmented channel of size $m+1$ with the $m+1$ sender having power ΔP. Then we have

$$\begin{aligned}\overline{C}_{n,FB}(P_1,P_2,\cdots,P_m) &\leq \overline{C}_{n,FB}(P_1,P_2,\cdots,P_m,\Delta P)\\ &\leq 2C_n(P_1,P_2,\cdots,P_m,\Delta P)\\ &\qquad\text{(result for even } m). \quad (46)\end{aligned}$$

Now by taking the limit $\Delta P \to 0$ in (46) we obtain the desired result for the channel with an odd number of senders:

$$\overline{C}_{n,FB}(P_1,P_2,\cdots,P_m) \leq 2C_n(P_1,P_2,\cdots,P_m).$$

Thus we have proved that feedback at most doubles the total capacity of an m-user additive Gaussian non-white noise multiple access channel.

VI. Conclusion

We have shown that the total capacity of any multiple access channel with non-white Gaussian noise can at most be doubled using feedback. Though we have not said much about achievability, one would suspect that there exists a generalization of the method described in [10] with the class of joint distributions induced by the linear feedback scheme $X_j^n = B_j Y^n + V_j^n$, $j = 1, 2, \cdots, m$, that may be used to characterize the capacity region. The general outer bound of Section III ((19)) might be relevant here. This technique may also provide insight in determining the capacity region of general m-user Gaussian white noise multiple access channels with feedback.

Appendix

Proof of Lemma 3.1:

$$
\begin{aligned}
&h(Y^n|W(S), W(\bar{S})) \\
&\overset{(i)}{=} \sum_{i=1}^{n} h(Y_i|W(S), W(\bar{S}), Y^{i-1}) \\
&\overset{(ii)}{=} \sum_{i=1}^{n} h\left(\sum_{j=1}^{m} X_{ij} + Z_i|W(S), W(\bar{S}), Y^{i-1}, \sum_{j=1}^{m} X_{ij}, Z^{i-1} \right) \\
&\overset{(iii)}{=} \sum_{i=1}^{n} h\left(Z_i|W(S), W(\bar{S}), Y^{i-1}, \sum_{j=1}^{m} X_{ij}\left(W_j, Y^{i-1}\right), Z^{i-1} \right) \\
&\overset{(iv)}{=} \sum_{i=1}^{n} h(Z_i|Z^{i-1}) \\
&\overset{(v)}{=} h(Z^n).
\end{aligned} \tag{A1}
$$

Here (i) is the chain rule, (ii) merely adds functions of the conditions, (iii) removes the conditionally deterministic constants X_{ij}, (iv) uses the conditional independence of $(W(S), W(\bar{S}), Y^{i-1}, \Sigma_{j=1}^{m} X_{ij})$ and Z_i given Z^{i-1}, and (v) unchains the chain rule. □

Proof of Theorem 3.1: Consider a sequence of

$$((2^{nR_1}, 2^{nR_2}, \cdots, 2^{nR_m}), n)$$

feedback codes

$$x_j^n\left(W_j, Y^{n-1}\right) = \left(x_{1j}(W_j), x_{2j}\left(W_j, Y^1\right), \cdots, x_{nj}\left(W_j, Y^{n-1}\right)\right), \quad j = 1, 2, \cdots, m.$$

Let $W_j \sim \text{unif}\{1, 2, \cdots, 2^{nR_j}\}$, $j = 1, 2, \cdots, m$, be mutually independent and also independent of the noise. We begin by using Fano's inequality to show that for any

$$((2^{nR_1}, 2^{nR_2}, \cdots, 2^{nR_m}), n)$$

code with $P_e^{(n)} \to 0$ and $\epsilon_n \to 0$ as $n \to \infty$, that

$$
\begin{aligned}
\sum_{j \in S} nR_j &= H(W(S)) \\
&\le I(W(S); Y^n) + n\epsilon_n \\
&= H(W(S)) - H(W(S)|Y^n) + n\epsilon_n \\
&\overset{(a)}{\le} H(W(S)|W(\bar{S})) - H(W(S)|Y^n, W(\bar{S})) + n\epsilon_n \\
&= I(W(S); Y^n|W(\bar{S})) + n\epsilon_n \\
&= h(Y^n|W(\bar{S})) - h(Y^n|W(S), W(\bar{S})) + n\epsilon_n \\
&\overset{(b)}{=} h(Y^n|W(\bar{S})) - h(Z^n) + n\epsilon_n \\
&\overset{(c)}{=} h\left(\sum_{j \in S} X_j^n - \sum_{j \in \bar{S}} X_j^n + Z^n \right) - h(Z^n) + n\epsilon_n.
\end{aligned} \tag{A2}
$$

Here (a) follows from the fact that $W(S)$ and $W(\bar{S})$ are independent and from the conditioning inequality $h(A|B) \ge h(A|B, C)$, (b) follows from Lemma 3.1, and (c) follows from Lemma 3.2 with

$$L_S = I^n, \qquad S \subseteq \{1, 2, \cdots, m\}.$$

If Z^n is multivariate Gaussian, (A3) follows from the expression for the entropy of a Gaussian vector and from the entropy maximizing property of the normal:

$$
\begin{aligned}
&h\left(\sum_{j \in S} X_j^n - \sum_{j \in \bar{S}} X_j^n + Z^n \right) - h(Z^n) + n\epsilon_n \\
&\qquad \le \frac{1}{2} \log \frac{\left| K^{(n)}_{\Sigma_{j \in S} X_j - \Sigma_{j \in \bar{S}} X_j + Z} \right|}{|K_Z^{(n)}|} + n\epsilon_n. \quad \square
\end{aligned} \tag{A3}
$$

References

[1] M. Pinsker, Talk delivered at the Soviet Information Theory Meeting, 1969 (no abstract published).

[2] P. Ebert, "The capacity of the Gaussian channel with feedback," *Bell Syst. Tech. J.*, pp. 1705–1712, October 1970.

[3] C. W. Keilers, "The capacity of the spectral Gaussian multiple-access channel," Ph.D. thesis, Dep. Elec. Eng., Stanford Univ., Stanford, CA, May 1976.

[4] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback—Part 1. No bandwidth constraints," *IEEE Trans. Inform. Theory*, vol. IT-12, pp. 172–182, Apr. 1966.

[5] N. T. Gaarder and J. K. Wolf, "The capacity region for a multiple access discrete memoryless channel can increase with feedback," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 100–102, Jan. 1975.

[6] T. M. Cover and C. S. K. Leung, "An achievable rate region for the multiple access channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 292–298, May 1981.

[7] F. M. J. Willems, "The feedback capacity of a class of discrete memoryless multiple access channels," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 93–95, Jan. 1982.

[8] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 623–629, July 1984.

[9] S. Ihara, "Capacity of discrete time Gaussian channel with and without feedback—I," *Mem. Fac. Sci. Kochi Univ. (Math.)*, vol. 9, pp. 21–36, March 1988.

[10] T. M. Cover and S. Pombra, "Gaussian feedback capacity," *IEEE Trans. Inform. Theory*, vol. 35, pp. 37–43, Jan. 1989.

[11] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inform. Theory*, vol. IT-33, pp. 711–716, Sept. 1987.

[12] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*. Boston: Allyn and Bacon, Inc., 1964.

[13] S. Pombra, "Non-white Gaussian channels with feedback," Ph.D. thesis, Dep. Elec. Eng., Stanford Univ., Stanford, CA, August 1991.

[14] S. Verdú, "Multiple-access channels with memory with and without frame synchronism," *IEEE Trans. Inform. Theory*, vol. 35, pp. 605–619, May 1989.

Optimal Shaping Properties of the Truncated Polydisc

Frank R. Kschischang, *Member, IEEE*, and Subbarayan Pasupathy, *Fellow, IEEE*

***Abstract*—Multidimensional constellation shaping with a family of regions called truncated polydiscs is studied. This family achieves maximum shaping gain for a given two-dimensional peak-to-average energy ratio or a given two-dimensional constellation expansion ratio. An**

Manuscript received April 24, 1992; revised April 6, 1993. This work was supported in part by the National Sciences and Engineering Research Council of Canada.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, Ont., Canada M5S 1A4.

IEEE Log Number 9401705.